Final Office Model Replication

This final guidance document presents all variable calculation and filter information, along with the final regression format, coefficients, and lookup table.

1. Notes on Variable Calculation

The final model requires a variety of calculations on the CBECS data in order to apply filters and put all of the variables in the right format. The following notes detail these calculations.

Energy Calculations

- Recodes: Observations that report "no use" for a given fuel type are not asked for a quantity. Therefore, for these observations it is appropriate to recode the system missing values to zero.
- Source Energy: The following factors are applied to convert values in CBECS into Source Energy
  - Electricity = 3.34
  - Natural Gas = 1.047
  - Fuel Oil = 1.01
  - Propane = 1.01
  - District Heat = 1.45 (if STUSED8=1)
  - District Heat = 1.35 (if only Hot Water is used, STUSED8=2)
- Propane: The amount of propane must be estimated from the category selected. This will involve a "maximum" estimation for the purpose of applying filters and an "actual" estimation for use in the rating models. For the maximum, the high end of the propane category is used. For the actual, the middle of the category is used. This is summarized in the table below.

Table 1
Propane Estimation Guidelines

PRAMTC8	Maximum Estimation (filter)	Actual Estimation (model)
1: less than 100	99	50
2: 100 to 499	499	300
3: 500 to 999	999	750
4,7,8,9: 1000 or higher, or unknown	Not Included	Not Included

In order to convert the amount of propane in gallons or pounds (indicated by PRUNIT8) into a Site energy value in Kbtu, the following conversions were applied:

- 1 Gallon of Propane = 92 Kbtu
- 1 Gallon of Propane = 4.5 Pounds

General Calculations

The following variables are necessary for the calculation of the model:

- Source Intensity: SourceEUI = SourceEnergy/Sqft8
- Natural Log of Square foot: LNSqFt = LN(SqFt8)
- Computer Density[1]: PCDen: 1000*PCNUM8/SqFt8
- Natural Log of Weekly Hours: LNWkHrs = LN(WKHRS8)
- Natural Log of Worker Density: LNWkrDen = LN(1000*NWker8/SqFt8)
- Commercial Refrigeration[2] Density:

RfgComDen = 1000*(RFGWIN8+RFGOPN8+RFGCLN8+RFGVNN8) / SqFt8

- Residential Refrigeration Density:

  RfgResDen = 1000*RFGRSN8/SqFt8
- Heating Degree Days x Percent Heated: HDDxPH = HDD658*HeatP8/100
- Cooling Degree Days x Percent Cooled: CDDxPC = CDD658*CoolP8/100

Banks

- For the purpose of the model a "Small Bank" is defined as a bank that is 50,000 square foot or smaller. This variable is used as a dummy variable in the model and also as an interactive term with both LnSqFt and LNWkrDen.

Variable Centering

For the purpose of the model, all of the variables are centered. Centering creates a situation where the intercept is equal to the average source energy intensity and the coefficients are used to adjust a building based on its deviation from the average value of each operating characteristic.

- Centered Variable = Observation Value – Weighted Mean for that Variable
- Centered LnSqFt = LnSqFt – Weighted Average LnSqFt
- Note that the weighted averages are computed across the filtered data set

A table of final centering values is attached (Attachment A).

2. Filters

Filters are established for a couple of reasons: to account for technical limitations of the CBECS data, to maintain baseline operational characteristics deemed appropriate for the program, and to prevent outliers from driving the regression result. The following table summarizes the filters for the Office model, alongside the rationale for each filter and the number of observations that remain after each filter. Once all of the filters have been applied, there are 498 observations remaining.

[1] All operating characteristic density values are computed on a per 1000 square foot basis

[2] To compute the refrigeration densities, note that if a type of refrigeration is not present, then the number for that type of units is not requested. In these situations the "System Missing" response can be re-coded to a value of zero.

Table 2 Summary of Office Model Filters		
Condition for *Including* an Observation in the Analysis	**Rationale**	**Number Remaining**
PBAPLUS8=2,3,4 or COURT8=1	All Offices, Banks, and Courts – Starting Set	808
Must have a weight (value for ADJWT8)	CBECS Limitation – Cannot incorporate un-weighted observations into analysis	755
Must have square foot <=1,000,000	CBECS Limitation – Actual values above 10^6 are not reported	728
If propane is used, the amount category (PRAMTC8) must equal 1, 2, or 3	CBECS Limitation – Cannot estimate propane use if it is "greater than 1000" or unknown	715
If propane is used, the maximum estimated propane amount must be 10% or less of the total source energy[3]	CBECS Limitation – Estimation of propane cannot introduce too much error into the energy use value	712
Must not use chilled water	CBECS Limitation – Quantities of chilled water are not collected	674
Must have at least 1 personal computer	EPA Requirement – To be considered a functioning office building, there must be PCs	669
Must operate for at least 30 hours per week	EPA Requirement – Baseline condition for being a full time office building	666
Must operate for at least 10 months per year	EPA Requirement – Baseline condition for being a full time office building	649
A single activity must characterize greater than 50% of the floor space[4]	EPA Requirement – In order to be considered part of the office peer group, more than 50% of the building must be Office, Bank, or Court space	625
Must have square foot >= 5,000	Analytical Limitation – Analysis could not model behavior for these smaller buildings	498

3. Regression

The final regression is a weighted ordinary least squares regression across the filtered data set of 498 observations. The dependent variable is SourceEUI. Each independent variable is centered as described in Section 1. The final model is estimated as follows.

[3] Here, the high end of the propane usage amount category (PRAMT8) is used to determine a maximum propane consumption and estimate what percent this would be of the total source energy. If that percent exceeds 10% of the total, then the observation is removed from the analysis.

[4] Generally, this is determined by a couple of screens. If the variable ONEACT8=1, this indicates that one activity is in 75% or more of the building. If the variable ONEACT8=2, then the variable ACT1PCT8 must be checked to determine what percent the main activity occupies, ACT1PCT8 must be greater than 50%. Analysis of the court observations shows that in some cases ACT2PCT8 is greater than ACT1PCT8. In these cases both activities are typical of courthouses, classified as either Office or Public Order. As such, the activity filter is not used to remove any court observations.

Table 3 Model Summary Output - SPSS				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.578(a)	.334	.319	2143.56856

a Predictors: (Constant), BANK_50, CCDDxPC, CLNWkHrs, CPCDen, CRfgComDen, CRfgResDen, CLnSqFt, Bank_50xCLNWkrDen, Bank_50xCLNSqFt, CHDDxPH, CLNWkrDen

Table 4 Model ANOVA Output - SPSS					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1121782970.818	11	101980270.074	22.194	.000(a)
Residual	2233114670.872	486	4594886.154		
Total	3354897641.690	497			

a Predictors: (Constant), BANK_50, CCDDxPC, CLNWkHrs, CPCDen, CRfgComDen, CRfgResDen, CLnSqFt, Bank_50xCLNWkrDen, Bank_50xCLNSqFt, CHDDxPH, CLNWkrDen
b Dependent Variable: SrcEUI
c Weighted Least Squares Regression - Weighted by ADJWT8 ADJWT8

Table 5 Model Coefficient Estimates Output - SPSS				
	Unstandardized Coefficients	**Standard Error**	**t**	**Sig.**
(Constant)	186.609	4.699	39.714	.000
CLnSqFt	34.175	5.271	6.484	.000
CPCDen	17.278	3.645	4.739	.000
CLNWkHrs	55.959	13.532	4.135	.000
CLNWkrDen	10.340	7.304	1.416	.157
CHDDxPH	.008	.003	2.962	.003
CCDDxPC	.014	.006	2.253	.025
CRfgComDen	277.698	57.101	4.863	.000
CRfgResDen	84.349	32.078	2.629	.009
Bank_50xCLNSqFt	-64.827	20.253	-3.201	.001
Bank_50xCLNWkrDen	34.202	15.884	2.153	.032
BANK_50	56.300	15.010	3.751	.000

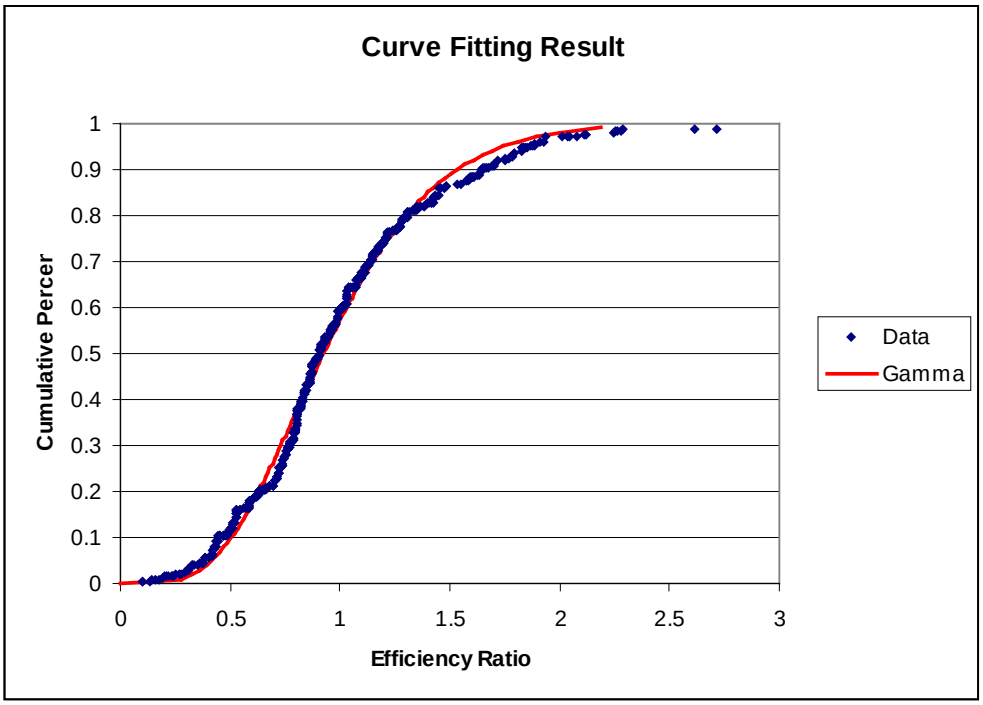
a Dependent Variable: SrcEUI
b Weighted Least Squares Regression - Weighted by ADJWT8 ADJWT8

4. Lookup Table

The model is used to generate a scoring lookup table according to the following steps:

- The model is used to generate a predicted Source EUI for each observation. Note that for programmatic reasons, each building is assigned the average value for Commercial Refrigeration Density and Residential Refrigeration Density. When each value is assigned the average, the centered value becomes zero.

- An energy efficiency ratio is calculated for each observation as follows:
  Actual Source EUI/Predicted Source EUI
- The weighted cumulative distribution of the energy efficiency ratios is fitted to a gamma distribution, where the parameters alpha and beta are estimated in order to minimize the sum of square differences between the actual and the gamma value for the cumulative distribution of each observation[5].
- For the office data the final gamma distribution is computed to have the following values:
  - Alpha = 5.64557
  - Beta = 0.17408
- The validity of this fit can be verified graphically



- The final gamma shape and scale parameters (alpha and beta, respectively) are used to calculate the efficiency ratio at each percentile (1 to 100) along the curve. These ratios are the lookup table. The final lookup table is attached (Attachment B).

[5] This fit can be achieved using MS Excel's Solver function. The actual cumulative distribution (weighted) is computed for each observation. Then, arbitrary alpha and beta values are assigned and used to compute a gamma distribution value for each observation. The solver can be set to minimize the sum of the differences between the actual and gamma distribution values across all observations, by changing both alpha and beta. The resulting alpha and beta are the final shape and scale values, respectively.

Attachment A – Table of Centering Values

Table of Centering Values		
Variable	**Full Name**	**Center Value (Weighted Mean)**
LNSqFt	Natural Log of Square foot	9.534603344
PCDen	PC Density	2.230828714
LNWkHrs	Natural Log of Weekly Hours	3.971610138
LNWkrDen	Natural Log of Worker Density	0.561566885
HDDxPH	Heating Degree Days x Percente Heated	4411.015712
CDDxPC	Cooling Degree Days x Percent Cooled	1156.570789
RFGComDen	Commercial Refrigeration Density	0.055462962
RFGResDen	Residential Refrigeration Density	0.131407891
Note: *Computed over the filtered data set (n=498 observations)* *Values are weighted by ADJWT8*		

Attachment B – Lookup Table

Rating	Cumulative Percent	Ratio
100	0	0
99	0.01	0.278705054
98	0.02	0.328378785
97	0.03	0.363069513
96	0.04	0.390860472
95	0.05	0.414570191
94	0.06	0.435548169
93	0.07	0.454555724
92	0.08	0.472068834
91	0.09	0.488406944
90	0.1	0.503795953
89	0.11	0.518402156
88	0.12	0.532351935
87	0.13	0.545743906
86	0.14	0.558656676
85	0.15	0.571154088
84	0.16	0.583288821
83	0.17	0.595104957
82	0.18	0.606639831
81	0.19	0.61792541
80	0.2	0.628989379
79	0.21	0.639855874
78	0.22	0.650546148
77	0.23	0.661079059
76	0.24	0.671471454
75	0.25	0.681738493
74	0.26	0.691893906
73	0.27	0.701950188
72	0.28	0.711918839
71	0.29	0.721810423
70	0.3	0.731634749
69	0.31	0.741400967
68	0.32	0.751117652
67	0.33	0.760792889
66	0.34	0.770434339
65	0.35	0.780049297
64	0.36	0.789644751
63	0.37	0.799227395
62	0.38	0.808803777
61	0.39	0.818380192
60	0.4	0.827962809
59	0.41	0.837557683
58	0.42	0.847170782
57	0.43	0.856808019
56	0.44	0.866475277
55	0.45	0.876178438
54	0.46	0.885923374
53	0.47	0.895716094
52	0.48	0.905562599
51	0.49	0.915469015

Rating	Cumulative Percent	Ratio
50	0.5	0.925441595
49	0.51	0.935486743
48	0.52	0.945611048
47	0.53	0.955821308
46	0.54	0.966124562
45	0.55	0.976528131
44	0.56	0.987039576
43	0.57	0.997667003
42	0.58	1.008418734
41	0.59	1.019303594
40	0.6	1.030330898
39	0.61	1.04151052
38	0.62	1.052852951
37	0.63	1.064369374
36	0.64	1.076071746
35	0.65	1.087972892
34	0.66	1.100086608
33	0.67	1.112427786
32	0.68	1.125012549
31	0.69	1.137858414
30	0.7	1.15098447
29	0.71	1.164411615
28	0.72	1.17816278
27	0.73	1.192263249
26	0.74	1.206741002
25	0.75	1.221627143
24	0.76	1.236956405
23	0.77	1.252767759
22	0.78	1.269105168
21	0.79	1.286018498
20	0.8	1.303564664
19	0.81	1.321809057
18	0.82	1.340827351
17	0.83	1.360707829
16	0.84	1.381554399
15	0.85	1.403490574
14	0.86	1.426664811
13	0.87	1.451257812
12	0.88	1.47749271
11	0.89	1.505649653
10	0.9	1.536087272
9	0.91	1.569275363
8	0.92	1.605846676
7	0.93	1.646683052
6	0.94	1.693067609
5	0.95	1.746975249
4	0.96	1.811687387
3	0.97	1.893296432
2	0.98	2.0053174
1	0.99	2.190161245